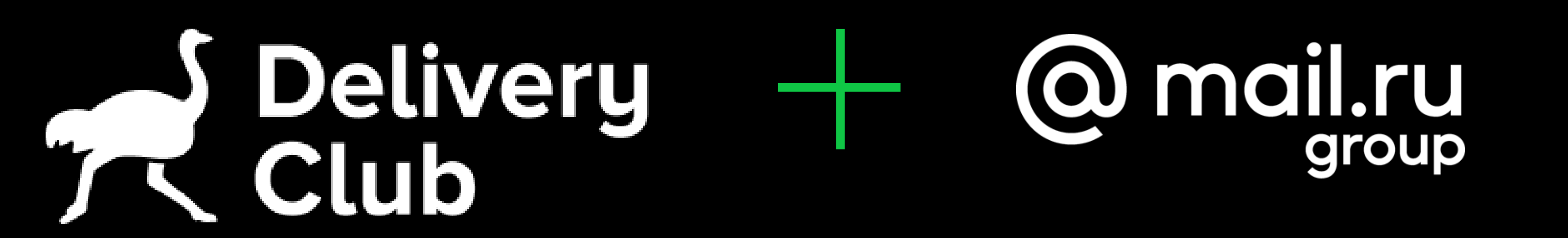
Геопространственное моделирование с применением методов машинного обучения

Измайлов Константин
Макин Евгений











Доставка еды №1 в России

Delivery Club входит в совместное предприятие со Сбербанком по развитию foodtech-направления

Бизнес-модель Delivery Club



Доставка заказов
собственными курьерами



Marketplace
Размещение предложения партнёра в приложении и на сайте сервиса

География доставки

Москва
Рязань
Нижний Новгород
Вологда
Курск
Тамбов
Казань

Екатеринбург
Тюмень
Омск
Ханты-Мансийск
Новосибирск
Хабаровск
Ростов-на-Дону

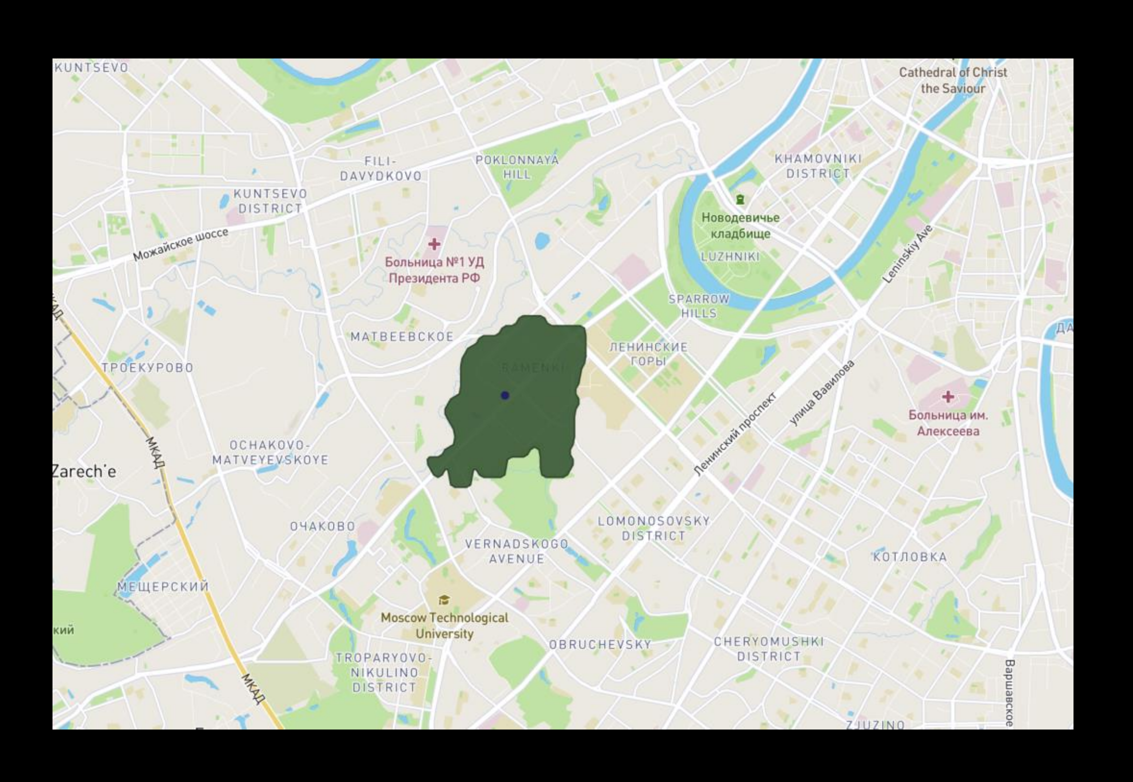
Краснодар
Пенза
Саратов
Архангельск
Воронеж
Оренбург

Более чем в 80 городах работают собственные курьеры

Баланс времени доставки и покрытия партнеров



Высокое время доставки
Высокое покрытие



Низкое время доставки
Низкое покрытие

Собственный «стиль» ручной отрисовки зон







Проблема ручного процесса

1

Непостоянство алгоритма, непрозрачная логика

2

Субъективизм в принятии решений о форме зоны и её характеристиках

3

Однообразный подход

4

Крайне низкая скорость

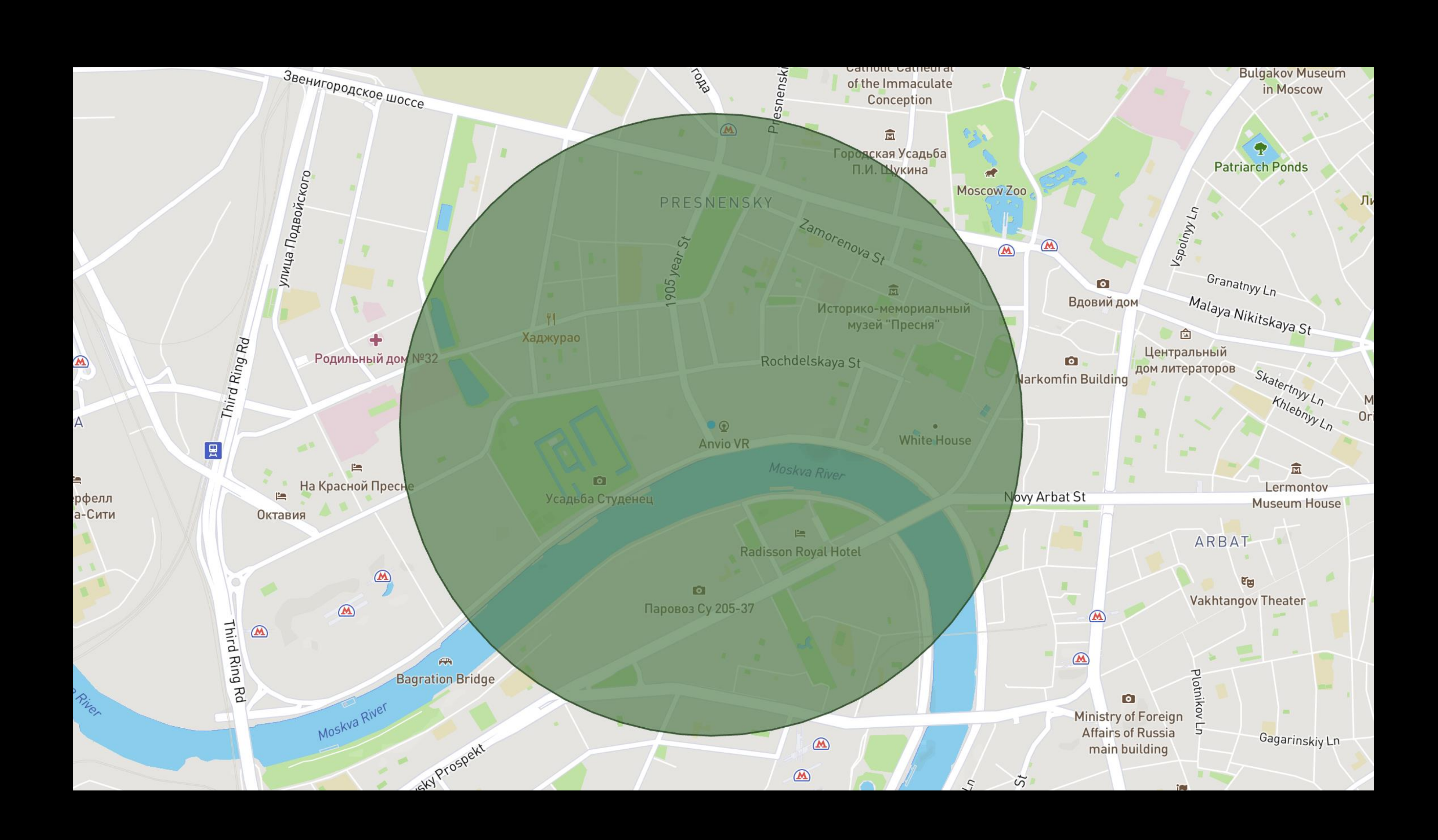
5

Ручной труд менеджеров по онбордингу ресторанов

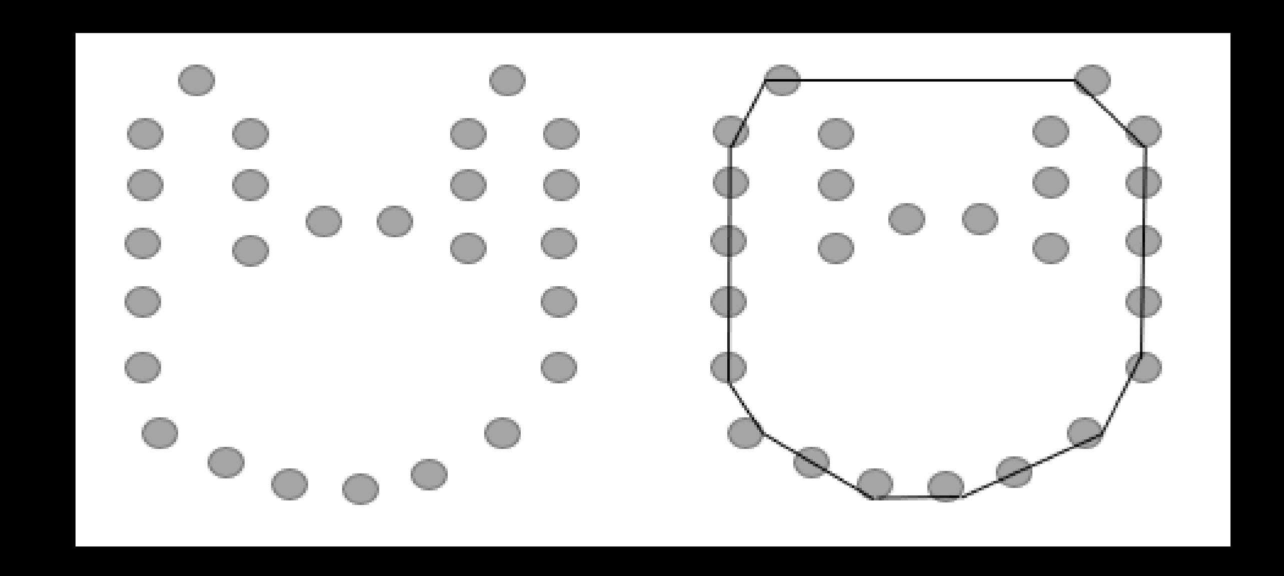
6

Отсутствие возможности регулирования цен доставки внутри зон ресторанов

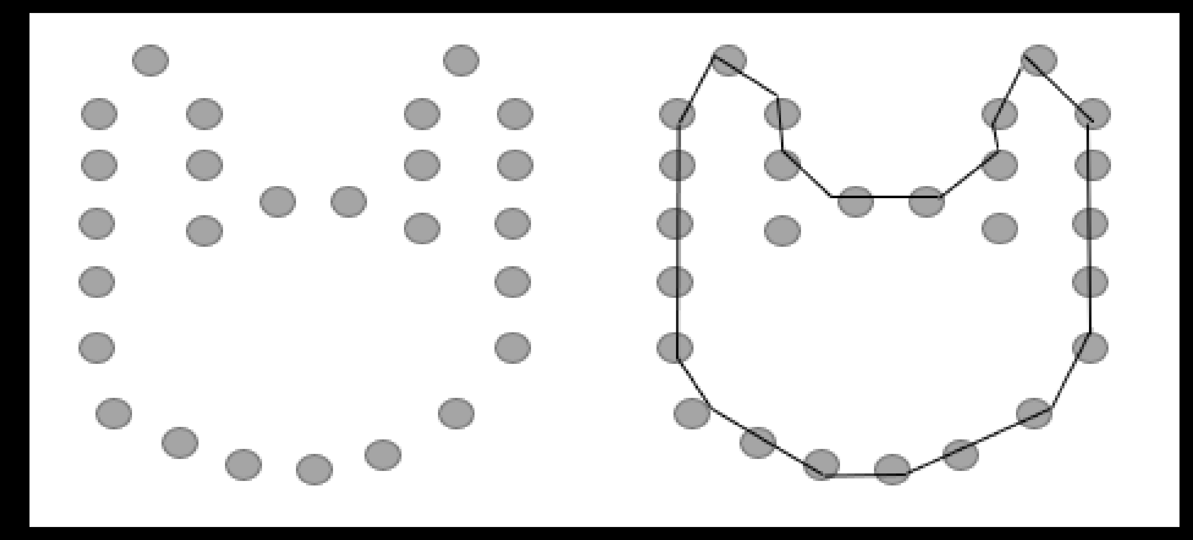
Простой круг



Усложнение алгоритма

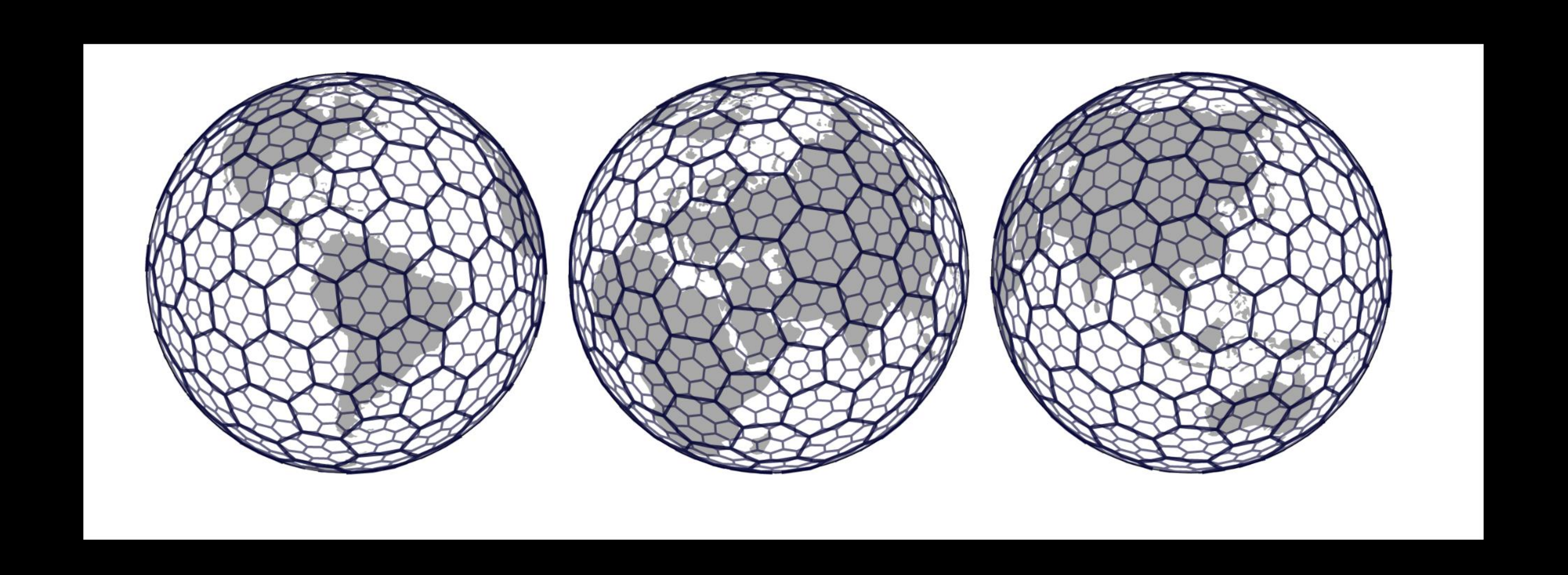


Convexhull



Alphashape

H3 — основа нашего алгоритма



Что это такое h3?



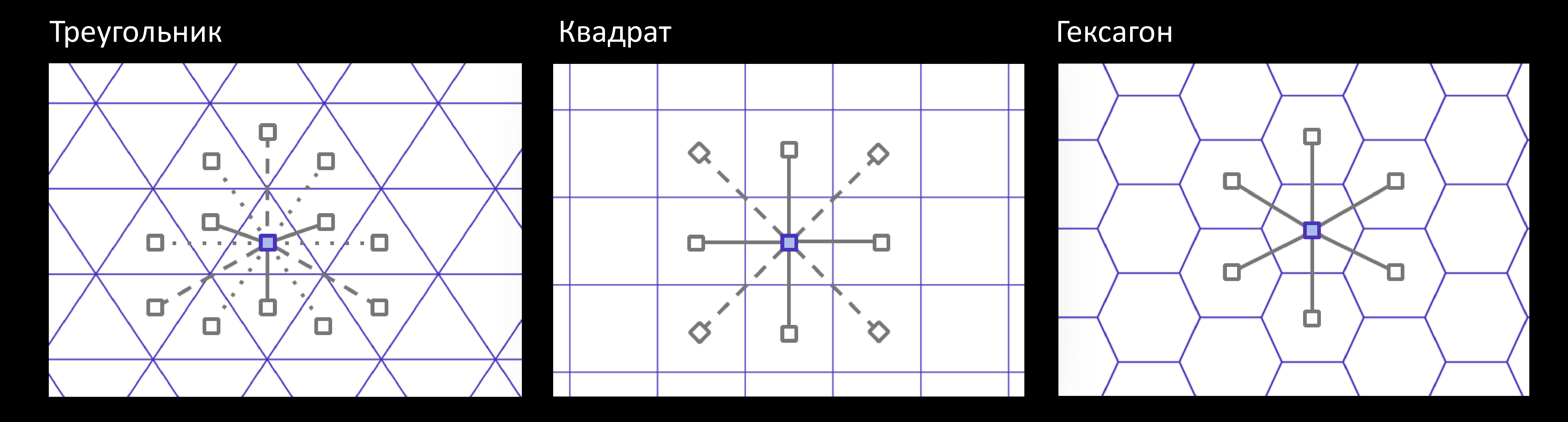


Почему мы выбрали h3?

Способы разбиения географической области на сегменты:

- Ручная разметка
- Регулярная сетка

На сегодня известно лишь три многоугольника, которые располагаются регулярно: треугольник, квадрат и гексагон. Значимым отличием между ними является расстояние до центров соседних ячеек.



Какие еще преимущества?

- ✓ Написана библиотека h3 Uber 
- ✓ Скорость вычислений
- ✓ Масштабируемость
- ✓ Возможность хранить в каждой ячейке любые свойства



Алгоритм создания зон



1

Мы создаем некий набор точек на поверхности вокруг ресторана и рассчитываем расстояние до каждой из них с помощью маршрутизатора

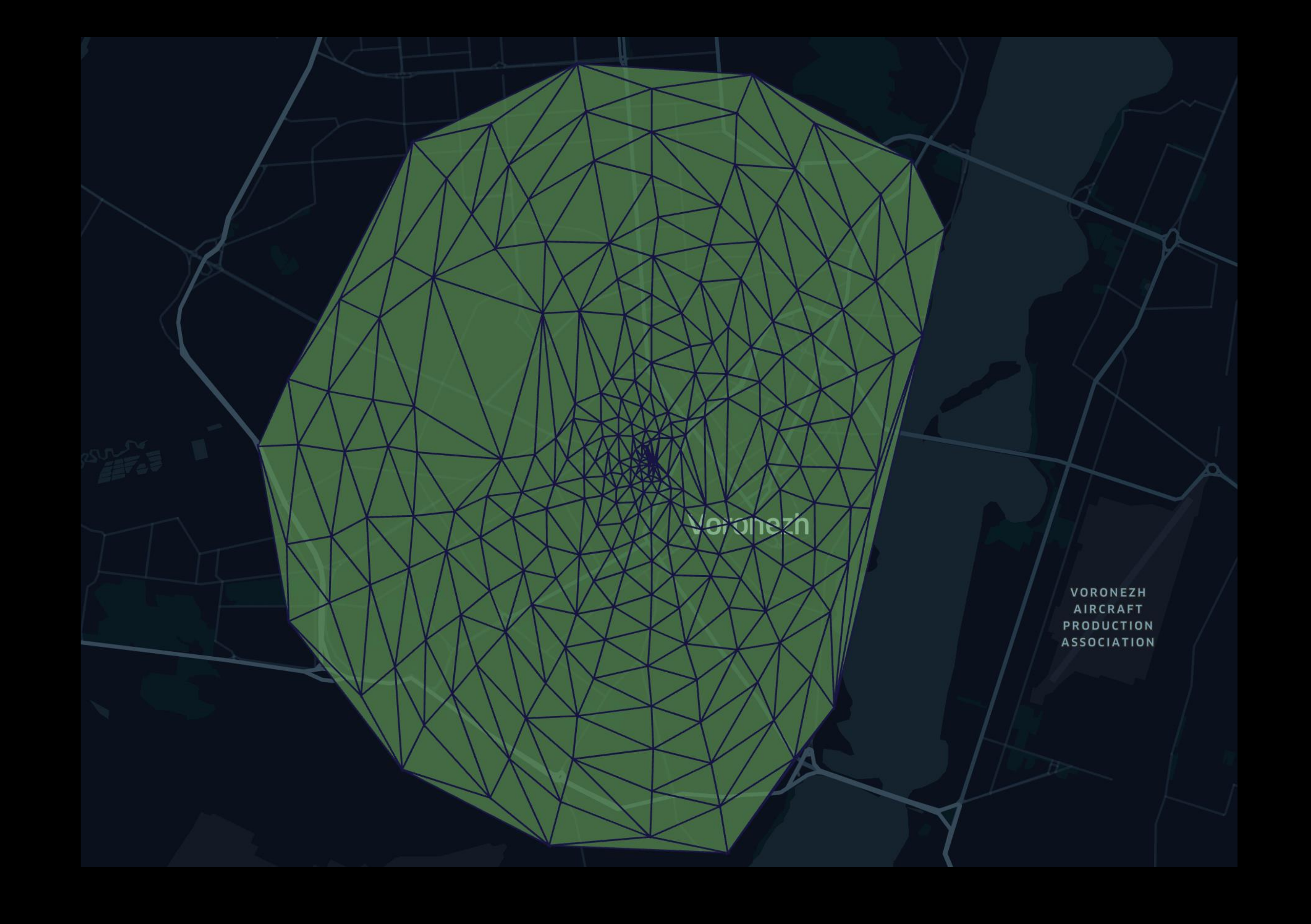
Voronezh

Voronezh

2 Далее, используя DBSCAN, мы убираем точки-выбросы



3 Исключаем некоторые точки из анализа, где ближайшая точка на графе находится слишком далеко



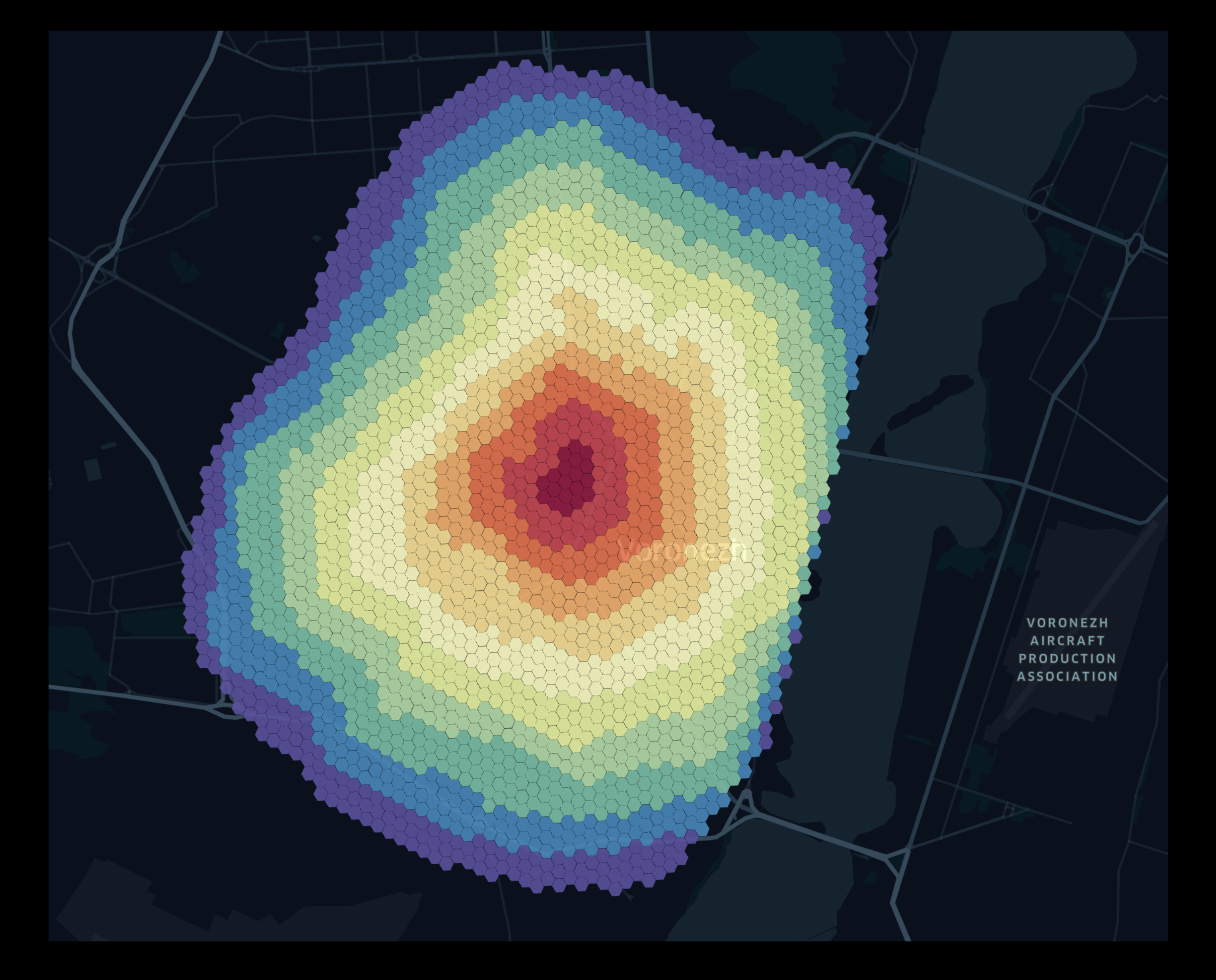
4

Далее на основе очищенных точек мы применяем триангуляцию Делоне



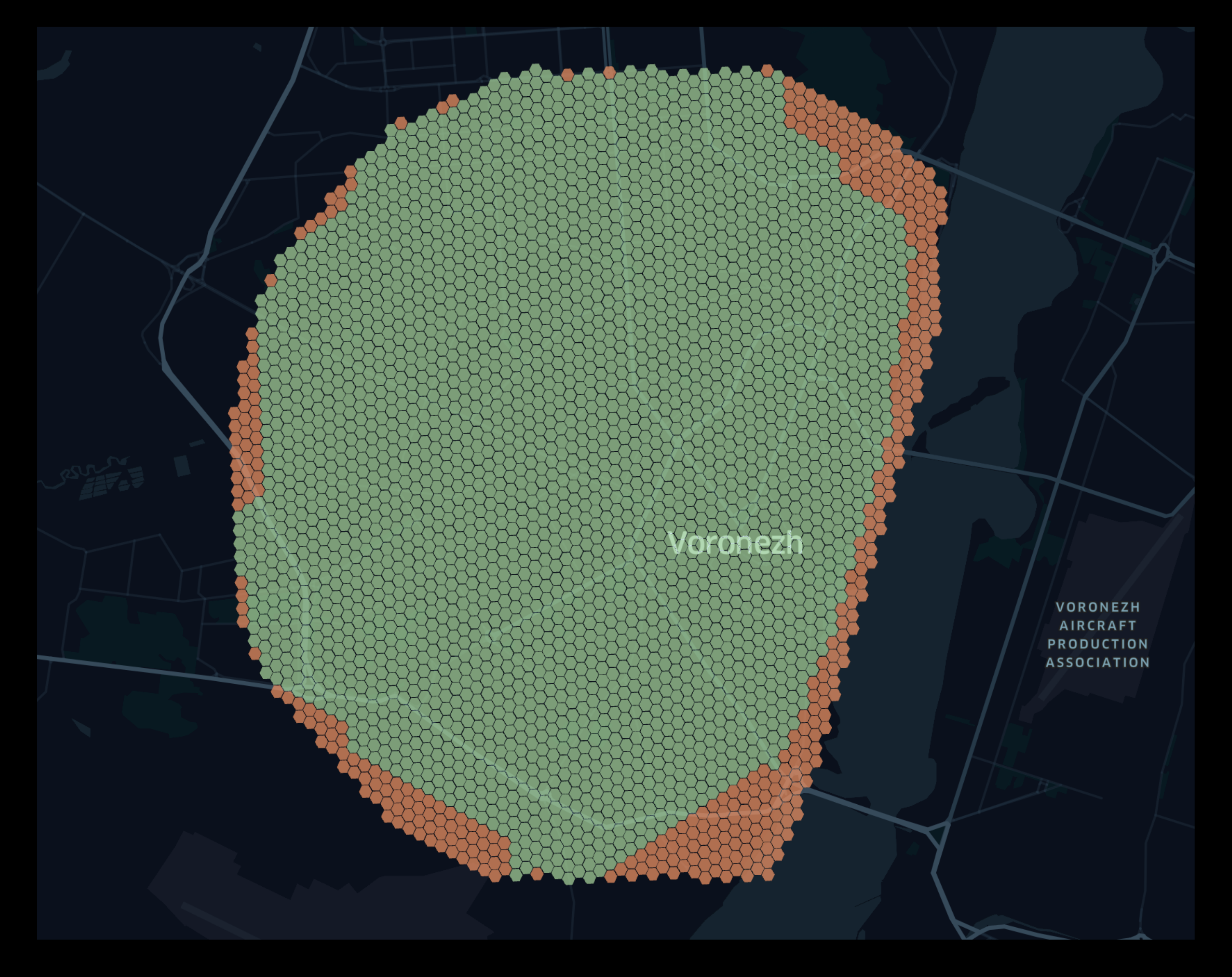
5

Далее мы создаем сетку гексагонов h3



6

Определяем, какому треугольнику принадлежит каждый гексагон h3, и затем для него определяем расстояние от ресторана (как аппроксимация по 3 вершинам треугольника)



7

Далее мы оптимизируем функцию ошибки, которая зависит от нашей задачи, постепенно *жадно* удаляя с внешней стороны зоны по одному гексагону

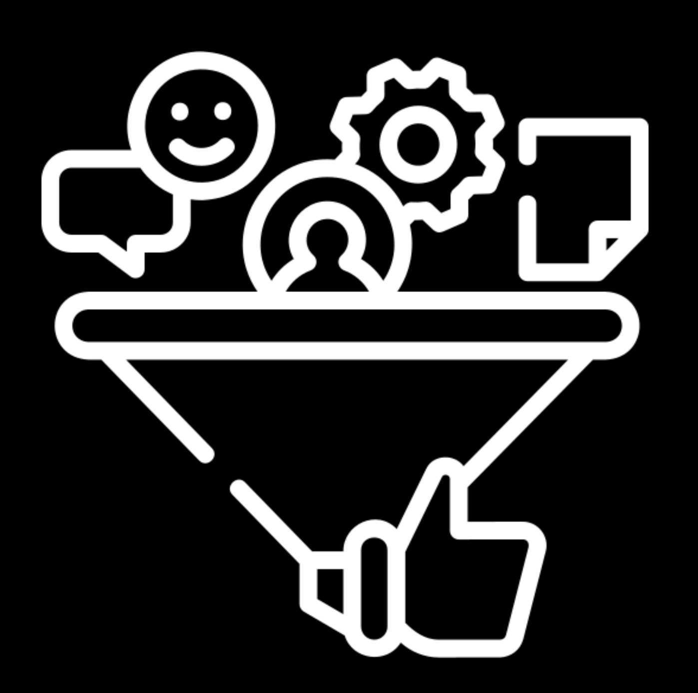


8

Последним шагом мы делаем временной градиент на получившихся зонах

Но как мы понимаем, какой гексагон нужно удалить?

Модель машинного обучения даёт прогноз для каждого гексагона зоны доставки каждого партнёра. Что использует модель:



Информацию о конкурентах ресторана

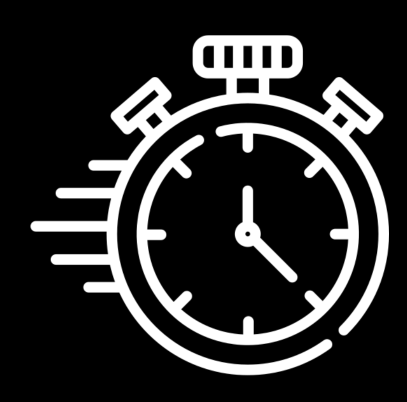
Население

Конверсию

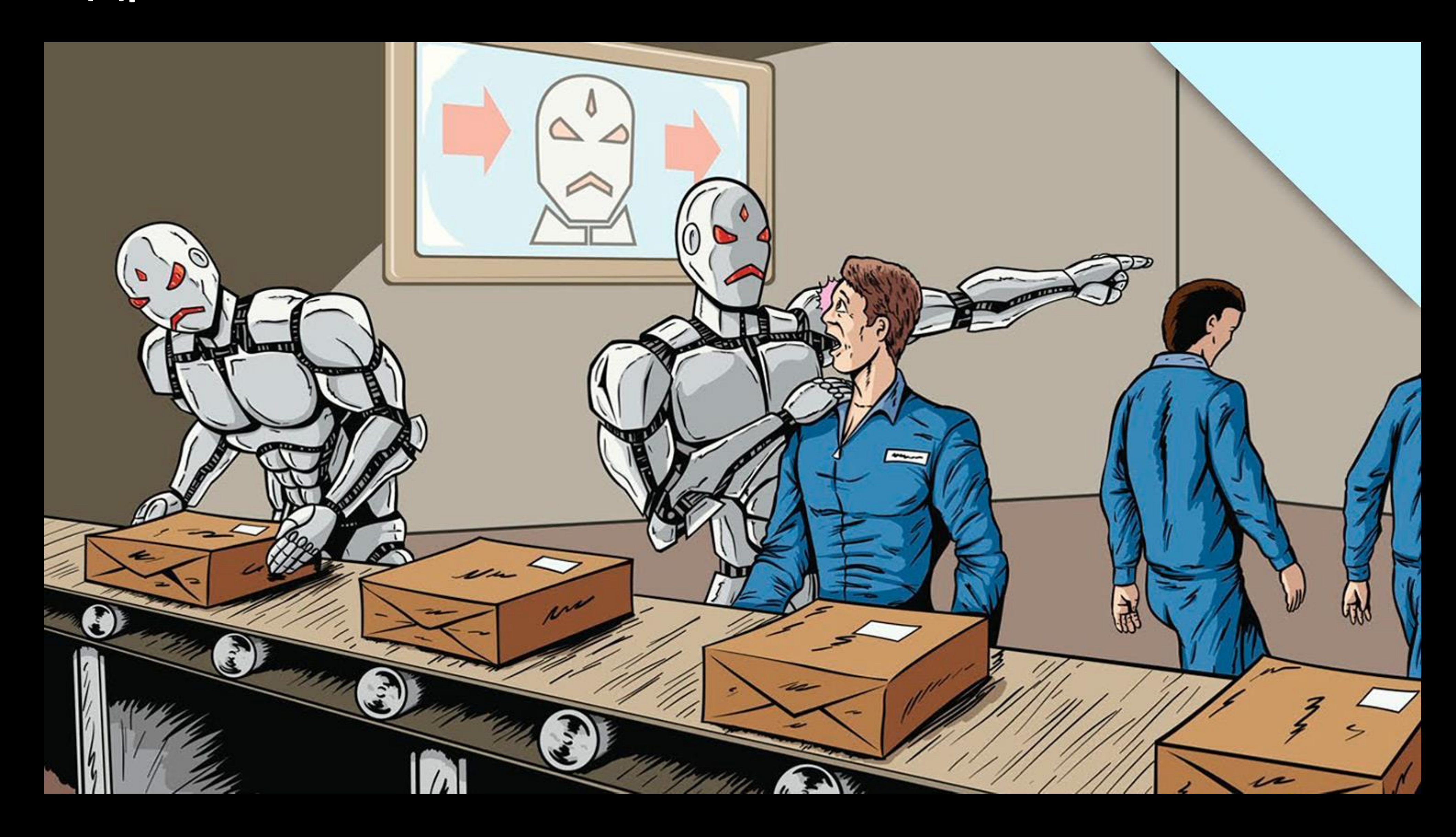
Информацию из внешних источников о популярности ресторана

Технические особенности реализации алгоритма, SLA?

Построение стандартной зоны доставки ресторана занимает на текущий момент ~ 1 сек



Внедрение



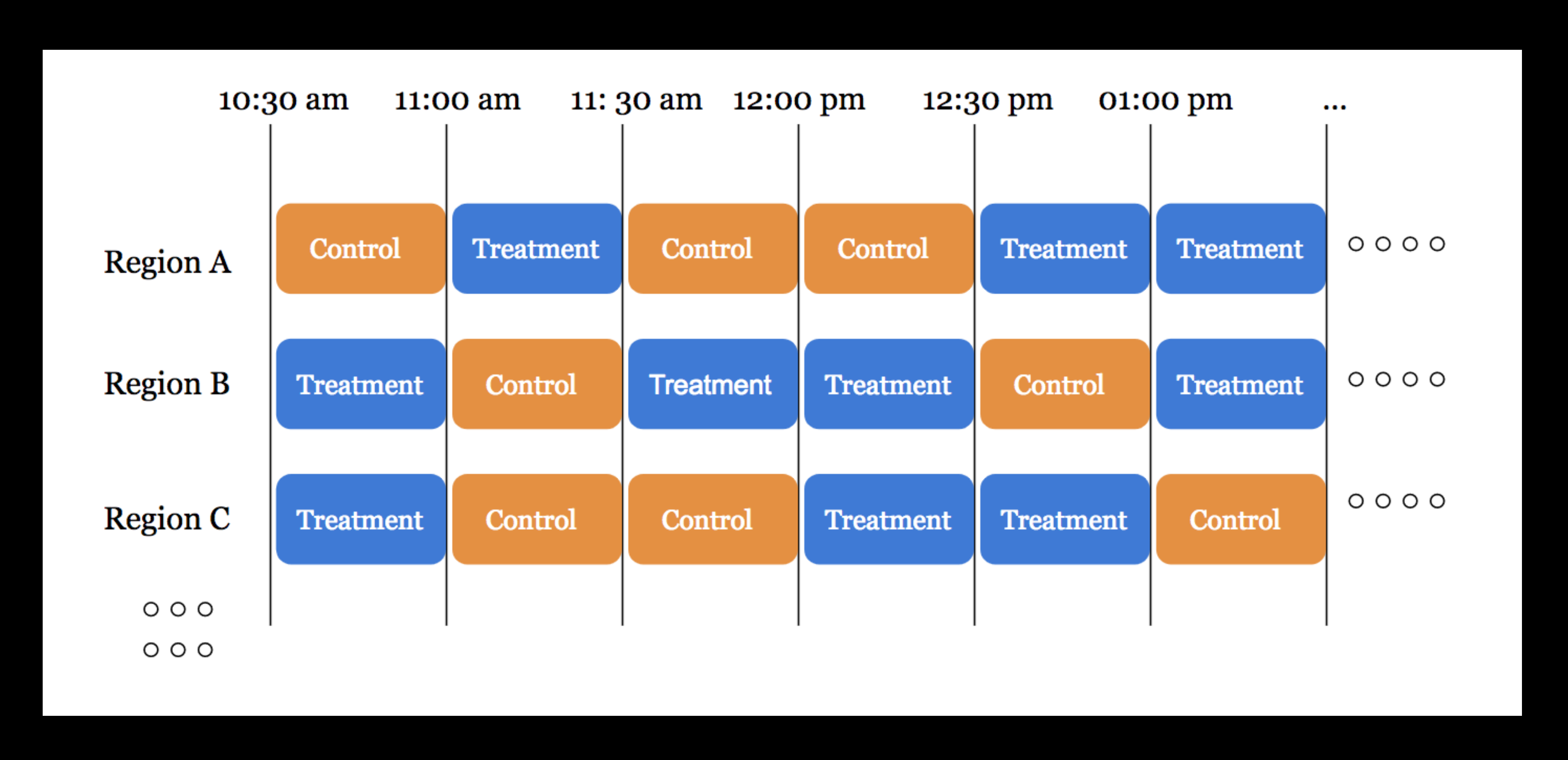
COVID comes…

- Тысячи ресторанов закрывают свои филиалы
- Увеличиваются зоны доставки ресторанов
- Сокращается меню и меняется целевая аудитория
- Меняется формат работы ресторанов (сотрудники становятся курьерами)
- Рост числа заказов



Оценка

SwitchBack A/B test



Что сейчас

- Сервис работает в продакшне. Зоны автоматически строятся по кнопке. При этом у нас появился функционал тестирования лесенки доставки

- 99% ресторанов перешли на алгоритмические зоны, и наш алгоритм поспособствовал переходу бэкенда на h3



В планах

- Новый функционал оптимизации
- Полная автоматизация процесса онбординга
- Ускорение алгоритма





Спасибо за внимание!